Identification and characterization of the constituent human serum antibodies elicited by vaccination

Jason J. Lavinder[a,b,1], Yariv Wine[a,b,1], Claudia Giesecke[c,d], Gregory C. Ippolito[e], Andrew P. Horton[f,g], Oana I. Lungu[a,b], Kam Hon Hoi[g], Brandon J. DeKosky[a], Ellen M. Murrin[a], Megan M. Wirth[a], Andrew D. Ellington[b,f,h,i], Thomas Dörner[c,d], Edward M. Marcotte[b,f,h,2], Daniel R. Boutz[b,f,2], and George Georgiou[a,b,e,g,2]

[a]Department of Chemical Engineering, [b]Institute for Cellular and Molecular Biology, [e]Section of Molecular Genetics and Microbiology, [f]Center for Systems and Synthetic Biology, [g]Department of Biomedical Engineering, [h]Department of Chemistry and Biochemistry, and [i]Applied Research Laboratories, University of Texas at Austin, Austin, TX 78712-1062; [c]Department of Medicine, Rheumatology, and Clinical Immunology, Charité University Medicine Berlin, 10098 Berlin, Germany; and [d]Deutsches Rheumaforschungszentrum Berlin, 10117 Berlin, Germany

Edited* by Stephen R. Quake, Stanford University, Stanford, CA, and approved January 2, 2014 (received for review September 24, 2013)

Most vaccines confer protection via the elicitation of serum antibodies, yet more than 100 y after the discovery of antibodies, the molecular composition of the human serum antibody repertoire to an antigen remains unknown. Using high-resolution liquid chromatography tandem MS proteomic analyses of serum antibodies coupled with next-generation sequencing of the V gene repertoire in peripheral B cells, we have delineated the human serum IgG and B-cell receptor repertoires following tetanus toxoid (TT) booster vaccination. We show that the TT^+ serum IgG repertoire comprises ∼100 antibody clonotypes, with three clonotypes accounting for >40% of the response. All 13 recombinant IgGs examined bound to vaccine antigen with $K_d \sim 10^{-8}$–10^{-10} M. Five of 13 IgGs recognized the same linear epitope on TT, occluding the binding site used by the toxin for cell entry, suggesting a possible explanation for the mechanism of protection conferred by the vaccine. Importantly, only a small fraction (<5%) of peripheral blood plasmablast clonotypes ($CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-TT^+$) at the peak of the response (day 7), and an even smaller fraction of memory B cells, were found to encode antibodies that could be detected in the serological memory response 9 mo postvaccination. This suggests that only a small fraction of responding peripheral B cells give rise to the bone marrow long-lived plasma cells responsible for the production of biologically relevant amounts of vaccine-specific antibodies (near or above the K_d). Collectively, our results reveal the nature and dynamics of the serological response to vaccination with direct implications for vaccine design and evaluation.

B-cell repertoire | proteomics

Most approved vaccines confer protection against infectious diseases by the induction of long-lived plasma cells (LLPCs), which secrete antibodies that serve to neutralize and opsonize the pathogen for many years or decades (1–3). Additionally, the generation of memory B cells (mBCs) provides both a mechanism for the rapid synthesis of affinity matured, antigen-specific antibodies following rechallenge and a means to diversify the humoral immune response to confer protection against rapidly evolving viruses or bacteria (4). Although some vaccines elicit antibody titers that remain virtually constant for many decades, for others, including the tetanus toxoid (TT) vaccine, antibody titers wane monotonically over time (5). Booster immunization triggers the rapid expansion and differentiation of cognate B cells, generating antigen-specific plasmablasts that peak in concentration in peripheral blood after 6–7 d and subsequently rapidly decline to nearly undetectable levels (6, 7). Some, but not all, of these peak-wave plasmablasts migrate to specialized niches overwhelmingly located in the bone marrow (BM) and survive as LLPCs (8), which constitute the major source of all classes of Ig in the serum (9).

The establishment of serological memory following either primary or booster vaccination is not understood well (10–14). Even though antibody production is the most critical effector function of B-cell immunity, and antigen-specific antibodies in the serum play a key role in protection against pathogen challenge, technical difficulties have precluded direct determination of the identities of the mAbs that comprise the serum antibody response to vaccination. However, recent studies showing that flu vaccination elicits not only neutralizing antibodies but also antibodies that enhance infection by different flu strains (15) underscore the pressing need to develop approaches for delineating the sequences and functionalities of the serum antibodies elicited by vaccination (16).

Single-cell cloning has been used to identify neutralizing antibodies encoded by mBCs or plasmablasts in peripheral blood (17). However, although extremely useful for understanding of the structural mechanisms that can lead to the blockade of pathogen infection, the interrogation of single peripheral B cells alone cannot provide information on whether antibodies encoded by single B cells are also produced as secreted IgGs from BM LLPCs, and hence whether they contribute to the serological memory induced by vaccination. A detailed understanding of the diversity of serum antibodies elicited by vaccination, their functionality (e.g., antigen affinity, epitope specificity), and their relative concentrations in the blood can provide key insights toward vaccine evaluation and development.

Here, we deployed high-resolution liquid chromatography (LC) tandem MS (MS/MS) (18–20) for the molecular-level analysis of the serum IgG repertoire, combined with deep sequencing of the V gene repertoire of peripheral B lymphocyte

Significance

Most vaccines confer immunity by eliciting long-term production of antibodies that bind to and neutralize the vaccine antigen. Remarkably, very little is known regarding the identities, sequence diversity, relative concentrations, or binding functionalities of the mAbs that comprise the serum repertoire elicited by vaccination. Here, we have delineated the constituent antibodies of the human serum IgG repertoire after vaccination and examined their relationship to the antibody V gene repertoire encoded by circulating B cells. The results detail the molecular composition and characteristics of the vaccine-specific serum antibody repertoire and demonstrate differences between the end-point response (the serum antibodies) and the peripheral B cells responding to the vaccine.

Author contributions: J.J.L., Y.W., C.G., T.D., and G.G. designed research; J.J.L., Y.W., C.G., G.C.I., A.P.H., O.I.L., K.H.H., B.J.D., E.M. Murrin, M.M.W., and D.R.B. performed research; J.J.L., Y.W., A.D.E., T.D., E.M. Marcotte, D.R.B., and G.G. analyzed data; and J.J.L., Y.W., C.G., G.C.I., T.D., D.R.B., and G.G. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]J.J.L. and Y.W. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: edward.marcotte@gmail.com, dboutz@mail.utexas.edu, or gg@che.utexas.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317793111/-/DCSupplemental.

subsets (20) and subsequent expression and characterization of representative serum antibodies, to map the dynamics of the human humoral response to vaccination in unprecedented detail. We elected to analyze the response to booster immunization of the TT vaccine because (*i*) it elicits a highly effective neutralizing response that is protective toward *Clostridium tetani* challenge; (*ii*) the vaccine is highly efficacious, and as a result, no deaths from tetanus intoxication have been reported in the United States for individuals who have completed at least primary immunization (21); (*iii*) TT has been used as a model for analyzing B-cell development following vaccination in humans (6, 22, 23); and (*iv*) although early serological and mAb studies had pointed to the C-terminal fragment of the toxin heavy chain [recombinant TT fragment C (rTT.C)] as the target for antibody-mediated protection (24), the precise mechanism by which antibodies elicited by the vaccine mediate neutralization has remained unclear.

We show that the anti-TT serum IgG repertoire at steady state is composed of a limited number of antibody clonotypes (~80–100) displaying uniformly high antigen affinity (low nanomolar or subnanomolar), that most of the serum repertoire postboost comprises preexisting (i.e., prevaccination) serum antibody clonotypes, and that there is only partial overlap between the peak-wave plasmablast V gene repertoire and the TT^+ serum IgG repertoire at steady state after vaccination. We identified several serum monoclonal IgGs that bind to rTT.C, and epitope mapping revealed that all rTT.C-specific antibodies tested bind to an immunodominant linear epitope at the ganglioside-binding site of the toxin that is used for cell entry. Computational antibody docking substantiated that binding of these antibodies to the toxin blocks access to the ganglioside ligand, thus providing a possible mechanistic explanation for how the TT vaccine confers protection. These results highlight the importance of understanding the composition and dynamics of the serum antibody repertoire, together with the V gene repertoire in peripheral B lymphocytes, for the molecular understanding of vaccine function.

Results

Sequencing of the Peripheral B-Cell Repertoire After TT Booster Vaccination. The inactivated TT vaccine is highly effective (21, 25), eliciting antibody titers in adults between 0.05 and 39.6 IU/mL, with a mean antibody titer of 1.2 IU/mL (20 μg/mL) (26). The TT-specific antibody titer decays linearly with a $t_{1/2}$ of 11 y (10); therefore, booster vaccinations are recommended every 10 y. Two healthy donors (HDs) were administered the TT/diphtheria toxoid (DT) vaccine 7 y and 10 y (HD1 and HD2, respectively) following the last booster vaccination. As expected, a rapid expansion of TT^+ plasmablasts ($CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-TT^+$) was observed at day 7 (Fig. S1). This wave of day 7 antigen-specific plasmablasts represents a transient population of responding B cells, with <30 TT^+ plasmablasts per milliliter detected at t = day 0 or at t = day 56 and beyond. The peripheral blood concentration of TT-specific mBCs remained relatively constant from t = 40 d to t = 169 d (Fig. S1). The V_H repertoires for each donor, encoded by day 7 plasmablasts and by IgD^- mBCs collected on both day 7 and 3 mo postboost, were determined by 454 (Roche Diagnostics GmbH) sequencing (70,326 and 157,089 high-quality V_H reads for HD1 and HD2, respectively; Table S1) and indexed by their V_H clonotype. The V_H clonotype, which represents a cluster of antibodies that likely originate from a single B-cell lineage (27, 28), is defined here as the group of V_H sequences that share germ-line V and J segments and also exhibit greater than 90% amino acid identity in the complementarity-determining region (CDR)-H3 (threshold for CDR-H3 amino acid identity determined by analysis of test sets from clustered deep-sequencing data; Fig. S2). We observed that the day 7 TT^+ plasmablast samples comprised 922 and 538 V_H clonotypes for HD1 and HD2, respectively.

Serum Proteomics of the TT-Specific IgG Repertoire. The TT^+ serum IgG repertoires at t = day 0, t = 7 d, t = 3 mo, and t = 9 mo postboost were analyzed using recently developed LC-MS/MS proteomic methodology (20). Importantly, in F(ab′)2 resulting from trypsin digestion of IgG, the presence of a conserved cleavage site (Arg) directly upstream of the CDR-H3 and at the fourth residue of the downstream CH1 constant region (Lys) consistently yields a peptide encompassing the highly informative CDR-H3 and the J region (Fig. S3). Proteolysis of the F(ab′)2 with other selective proteases (e.g., GluC/LysC) resulted in peptide identifications of very few additional clonotypes (<8% additional high-confidence identifications of those found in trypsinized sample for HD2 at day 0), the vast majority of which were of low abundance. For peptide identifications, a custom database of the antibody repertoire was built using high-quality V gene sequences from the peripheral B cells in each donor (Table S1), in conjunction with a standard shotgun proteomic pipeline with a high-mass accuracy filter (average mass deviation <1.5 ppm) to minimize false identifications (20). Frequencies of antigen affinity chromatography elution- and flow-through–derived CDR-H3 peptides mapping to a unique clonotype in the 454 donor-specific sequence database are shown in Fig. 1. The serum IgG clonotype frequency histograms are highly reproducible among technical replicates (20).

Sensitivity and Resolution of CDR-H3 Peptide Quantitation. To determine the dynamic range of detection of serum antibodies and to calibrate the resolution of antibody quantitation, isotopically labeled peptides corresponding to seven TT-specific CDR-H3 sequences observed over a wide range of MS peak intensities in serum samples from donor HD1 and ranging from 15 to 25 residues in length (i.e., largely spanning the observed CDR-H3 peptide length distribution) were synthesized and spiked into trypsinized HD1 samples at varying amounts (5–500 fmol). For all seven synthetic peptides, peak intensities varied linearly with peptide concentration (Spearman correlation = 0.98) and displayed small differences (less than threefold) across different peptides at each spike-in concentration (Fig. S4). The LC-MS/MS detection limit was found to be 5 fmol. Thus, based on the amount of trypsinized F(ab′)2 injected, we estimate the lower limit of sensitivity of IgG in the serum at ~0.1 nM (or ~15–16 ng/mL).

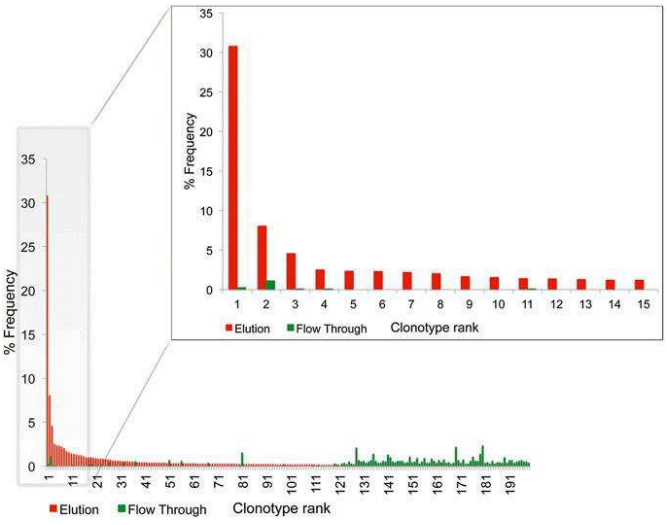

Fig. 1. Representative histogram of antibody clonotype frequencies identified proteomically in the F(ab′)2 elution and flow-through fractions following TT affinity purification. The histogram shown depicts the 3-mo postboost serum IgG repertoire for HD1. Frequencies shown here were calculated by adding the CDR-H3 spectral counts for all peptides mapping to a single clonotype.

Identities and Dynamics of the Serum Antibody Response to Vaccination. The composition, persistence, and dynamics of V_H clonotype frequencies in the TT-specific serum IgG repertoire at t = day 0, t = day 7, t = 3 mo, and t = 9 mo postboost are shown in Fig. 2; the t = day 0 and t = 9 mo time points constitute the steady-state response pre- and postboost vaccination. At steady state pre- and postboost, the TT^+ serum repertoire displays a comparable clonal diversity in both donors, comprising between 82 and 124 antibody clonotypes. Particularly striking features of the serological repertoire include the following:

i) For both donors at 9 mo postboost, >40% of the detectable serum antibody response (the sum of all TT^+ CDR-H3 peptide peak intensities) could be traced to three antibody clonotypes. The vast majority of the rest of the TT-specific clonotypes were present at frequencies under 0.5%. Specifically, in both donors, we estimate that the most abundant TT-specific antibody clonotype is present in the serum at a concentration of several micrograms per milliliter (~15–20 μg/mL based on the MS peptide peak intensity and peptide concentration calibration in Fig. S4).

ii) In both HD1 and HD2, highly abundant serum antibody clonotypes at day 0 were also present at high concentrations 9 mo postboost.

iii) A transient increase in the TT^+ serum IgG clonotypic diversity index (IgG-cd index) was observed at 3 mo postvaccination (82% increase in the IgG cd-index for HD1 and 17% increase for HD2); the smaller increase in the IgG-cd index for HD2 may be related to the fact that this donor did not display a higher TT titer postvaccination (Table S2) even though a robust cellular response to the vaccine was evident at day 7 (Fig. S1).

iv) The postboost steady-state repertoire comprised 59% (HD1) and 21% (HD2) new TT^+ antibody clonotypes not observed in the t = day 0 sample. The overwhelming majority of these were detected at low levels, with only one to two new antibody clonotypes in each donor detected at levels ≥0.5% of the antigen-specific serum repertoire.

Comparison of the Peripheral B-Cell and TT^+ Serum Antibody Clonotypes. Earlier studies in mice had revealed a wide disparity in clonal diversity between the antigen-specific peripheral B cells and the steady-state LLPC population in the BM that is predominantly responsible for the synthesis of the antibodies that constitute the serological memory (29). It had been estimated that only ~10% of responding peak plasmablasts following TT boost vaccination in humans differentiate further into LLPCs (30). Consistent with these observations, comparison of the steady-state TT^+ serum IgG repertoire with the repertoire encoded by the day 7 TT^+ plasmablasts revealed that <5% of the peak-wave (day 7) TT-specific plasmablast clonotypes contribute to the serological memory at a level sufficiently high enough to be detected (Fig. 3 and Table S1). Additionally, a small percentage (<1%) of the day 7 plasmablast clonotypes that did not stain with fluorescently labeled TT [i.e., TT depleted (TTdep)] encode antibodies that were detected in the TT^+ serum IgG repertoire. These TT^{dep} plasmablasts likely expressed insufficient levels of membrane IgG to stain with TT during fluorescence-activated cell sorting (FACS) yet encoded TT-specific antibodies, and consequently ended up contributing to the serological memory.

Many highly expanded clonotypes encoded by the day 7 TT^+ plasmablasts (as determined by high 454 sequencing read frequencies) were also detected at high or moderately high levels in the serum at 9 mo (Fig. 3). This indicates, albeit not directly, that these respective TT^+ plasmablasts likely developed into LLPCs at steady state after vaccination. Direct observation of these LLPCs over the course of a vaccine response is extremely difficult in humans, because most LLPCs reside in the BM. In both donors, there were, however, several highly expanded day 7 TT^+ plasmablast clonotypes that encoded antibodies not detected in the 9-mo serological repertoire. To examine whether the absence of these antibodies in the serum was due to a technical inability to detect the respective tryptic peptide fragments by MS, we selected one of the antibodies found at a very high 454 sequencing read frequency (9%) in the day 7 TT^+ plasmablast repertoire from HD2, yet absent from the serological repertoire at all time points, and expressed it in HEK 293F cells. We found that the CDR-H3 tryptic peptides of this recombinant antibody were readily observable by LC-MS/MS (Fig. S5). This result strongly suggests that the TT^+ plasmablast clonotypes that could not be detected by our proteomic analysis are either not present in the serum at all or are present at subphysiological concentrations, below the 0.1 nM LC-MS/MS detection limit.

Characterization of TT-Specific Antibodies. V_H/V_L pairs encoded by TT^+ plasmablasts from HD2 were determined by a high-throughput method (19), and nine of those for which the clonotype was detected in the TT^+ serum IgG repertoire were expressed as human IgG1 antibodies in HEK 293F cells. These experiments used previously frozen plasma cells, which are known to display low viability upon thawing; consequently, some V_H/V_L pairs of particular interest could not be determined. Therefore, for the highest frequency TT^+ serum clonotype from HD2 and three high-frequency TT^+ serum clonotypes from

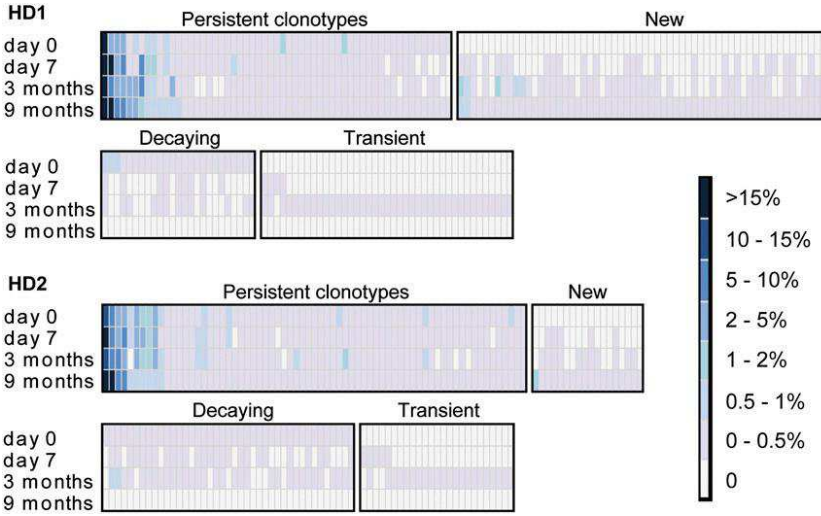

Fig. 2. Heat map of TT^+ serum IgG clonotypes. Binned frequencies of TT^+ serum clonotypes are determined from mass spectral peak intensities [extracted-ion chromatogram (XIC) peak areas] of CDR-H3 peptides. Each column represents a unique clonotype across four different time points relative to the booster vaccination. Persistent clonotypes (present at day 0 and 9 mo) and new clonotypes (not detected at day 0 yet detected at 9 mo) are ordered by the postboost steady-state (9 mo) frequency. Decaying clonotypes (present at day 0 but not detected at 9 mo) are ordered by the preboost steady-state (day 0) frequency. Transient clonotypes are not detected at pre- or postboost steady state.

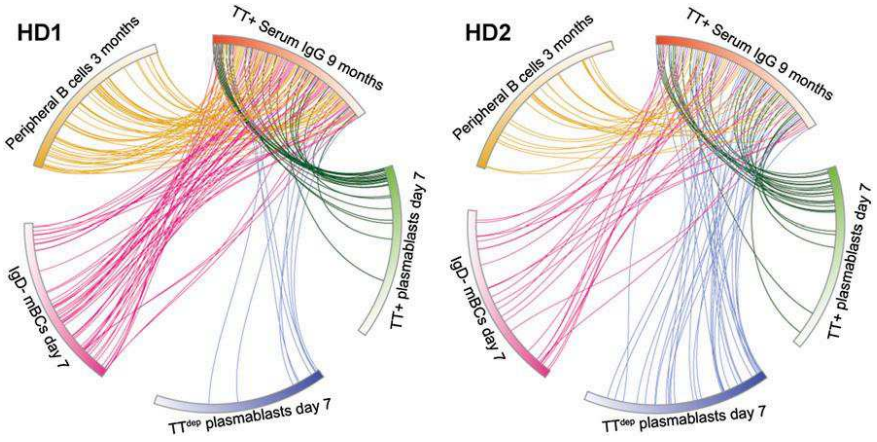

Fig. 3. Circos plots show the relationship between the B-cell and TT^+ serum IgG repertoire. Postboost steady-state (9 mo) TT^+ serum IgG clonotypes (red section, ordered clockwise by frequency) and the V gene repertoire clonotypes determined by 454 sequencing of each B-cell subset, ordered clockwise by frequency, in HD1 (*Left*) and HD2 (*Right*) are shown. IgD^- mBC day 7: $CD3^-CD14^-CD19^+CD27^+CD20^+IgD^-$, TT^+ plasmablasts day 7: $CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-TT^+$, TT^{dep} plasmablasts day 7: $CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-$. Even though some low-abundance TT^+ serum IgG clonotypes solely map to day 7 TT^{dep} plasmablasts, these cells did not stain with antigen by FACS, likely due to varying levels of B-cell receptor expression and affinities for the antigen-dye conjugate. Peripheral B cells 3 mo: combined IgD^- mBCs and total plasmablast cells. Clonotype numbers for each population are provided in Table S1.

HD1, V_H genes were first synthesized and used to construct four separate Fab libraries in bacteria with V_L cDNA from donor-specific TT^+ plasmablasts (20). Because of the low diversity of the TT^+ plasmablast-derived V_L pool (1,100–1,300 unique CDR-L3 genes), these libraries could be readily screened by ELISA, resulting in the isolation of productive V_H/V_L pairs for each clonotype (Fig. S6). Overall, we determined the affinity (K_d) of 13 full-length serum IgGs to vaccine-grade TT (formaldehyde-inactivated), to native rTT.C, and to formaldehyde-treated rTT.C (Fig. 4). As expected for booster vaccination, all antibodies bound the vaccine antigen TT with high affinities ($K_d \sim 2.2 \times 10^{-8} - 1 \times 10^{-10}$ M). Neutralizing epitopes of the 150-kDa holotoxin have been reported to reside within rTT.C (24). Five serum IgGs (HD1-2, HD1-5, HD2-49, HD2-88, and HD2-89) were shown to bind to rTT.C with subnanomolar K_d values. Epitope binding analysis using an overlapping 15-mer peptide array spanning the entire rTT.C polypeptide (110 peptides in the array) revealed that all five rTT.C-specific antibodies recognized a single linear epitope (Fig. 5). Modeling of the Ab/Ag complex of HD2-49, HD2-88, and HD2-89 (Fig. S7) using Rosetta software showed that binding of these antibodies to rTT.C occludes the ganglioside-binding site, which serves as the cell surface ligand used by the toxin for its internalization, and thus may play a critical role in toxin neutralization. The structural models reveal numerous contacts between the antibody CDR-H3 and CDR-L3 loops and the experimentally defined binding epitope of the toxin.

mAb ID	CDR-H3 sequence	TT^+ plasmablasts day 7 (%)	TT^+ serum IgG day 0	TT^+ serum IgG day 7	TT^+ serum IgG 3 months	TT^+ serum IgG 9 months	α-TT (inactivated) K_D(nM)	α-rTT.C (native) K_D(nM)	α-f-rTT.C (inactivated) K_D(nM)
HD1-1*	ARNLQGHYAMDV	0.03					0.3 ± 0.06	n.d.	n.d.
HD1-2*	ARDSVTNLGENLNFFPY	15					0.15 ± 0.03	1.67 ± 0.08	0.40 ± 0.07
HD1-5*	ARDTVTPLGENLNYFAH	1					0.55 ± 0.03	1.33 ± 0.12	0.20 ± 0.06
HD2-1*	ARGEPIRATVFGQPIPRGAWFDP	25					0.56 ± 0.03	n.d.	n.d.
HD2-7	ARLHPTCASTRCPENYGMDV	0.4					18.1 ± 3.6	n.d.	n.d.
HD2-8	ARARNYGFPHFFDF	0.04					0.5 ± 0.01	n.d.	n.d.
HD2-21	ARGEDCVGGSCYSAD	0.8					0.1 ± 0.008	n.d.	n.d.
HD2-26	ARDYFHSGSQYFFDY	1					0.6 ± 0.03	n.d.	n.d.
HD2-49	ARGVVPAGIPFDF	0.05					3.2 ± 0.5	0.31 ± 0.03	3.12 ± 0.32
HD2-54.1	AKAPIIGPKYYFYMDV	0.2					1.6 ± 0.04	n.d.	n.d.
HD2-54.2	AKAPIIGPKHYFYMDV	0.2					0.9 ± 0.03	n.d.	n.d.
HD2-88	ARKGMGHYFDF	0.9					2.8 ± 0.3	0.7 ± 0.13	2.50 ± 0.27
HD2-89	GKSYDYIRENLDS	0.3					22.6 ± 9	0.44 ± 0.03	8.21 ± 0.91
HD2-PB1	AKDRVRVVQAATTLDF	9					3.7 ± 0.5	n.d.	n.d.
HD2-PB2	ARKPRFYYDTSAWFEF	0.2					1.6 ± 0.05	n.d.	n.d.

*pairing of V_H:V_L derived from 96-well ELISA screening of V_L shuffle libraries

Fig. 4. Functional analysis of recombinant TT-specific mAbs from serum. V_H/V_L pairs from day 7 TT^+ plasmablasts were determined (19), genes were synthesized, and IgG antibodies were expressed in HEK 293F cells. For the four IgGs marked with an asterisk, the V_L genes were identified by screening libraries of each of the respective synthetic V_H genes paired with TT^+ plasmablast V_L repertoires on 96-well ELISA plates. The mAb ID indicates the donor from which the antibodies were derived, followed by the serum IgG clonotype ranking (by postboost steady-state frequency). For each V_H, the respective clonotype frequency (%) in the day 7 TT^+ plasmablast V gene repertoire is shown. The heat map indicates the relative frequency of the clonotype for each V_H (a heat map key is provided in Fig. 2). Equilibrium dissociation constants (K_d values) toward vaccine-grade TT (TT), native rTT.C, and formaldehyde-inactivated rTT.C (f-rTT.C) were determined by competitive ELISA. HD2-PB1 and HD2-PB2 correspond to V_H genes found in the day 7 TT^+ plasmablast V_H gene repertoire but not detected in the serum at any time point examined. MS analysis of trypsin digests of recombinant HD2-88 and HD2-PB1 is illustrated in Fig. S5. Full amino acid sequences of analyzed TT-specific mAbs are provided in Fig. S6. n.d., not determined.

Discussion

The majority of approved and experimental vaccines elicit long-lasting humoral immunity and a protective response that manifests as a significant titer of neutralizing serum antibodies (2). Vaccine development and evaluation have so far relied on the determination of neutralization titers, with titers higher than a certain threshold deemed protective. However, antibody titers constitute an aggregate property from which it is very difficult to infer, let alone to ascertain, the precise mechanism of protection conferred by the multiplicity of antibodies contained in the serum (31). Additionally, many experimental vaccines fail because they cannot elicit prolonged neutralizing antibody responses or because the vaccine response is directed toward nonneutralizing or heterologous epitopes (15); thus, it is important to be able to trace the identities, dynamics, and binding functionality of the monoclonal serum antibodies induced by vaccination to identify B cells encoding these antibodies, and to delineate their developmental trajectory and fate.

We have developed a platform technology for determining both the serological antibody repertoire and the V gene repertoires in peripheral B cells. Combined with expression of representative antibodies found in serum, antigen affinity and epitope mapping studies, and Rosetta modeling of antibody–antigen complexes, we show here that this pipeline can provide completely unique insights on the nature of humoral responses elicited by vaccination through a direct proxy of the serum antibody vaccine response. In two donors, we find that the anti-TT polyclonal response comprised approximately the same number of distinct antibody clonotypes at steady state (IgG-cd index = 80–120). As expected, a transient increase in the serological clonotypic diversity was observed after the peak response and was most pronounced at 3 mo after vaccination. For comparison, the serum of hyperimmunized rabbits exhibited a more restricted diversity of antigen-specific antibodies (IgG-cd index = 30) (20).

We note that the estimation of the IgG-cd index is subject to the established limitations of LC-MS/MS proteomic analysis and likely represents a lower estimate of the true diversity of the serological repertoire, because some CDR-H3–derived peptides are present below the level of detection. Nevertheless, several lines of evidence indicate that the LC-MS/MS serum proteomics we used here clearly capture the salient features of the serological repertoire. First, digestion of serum antibodies with proteases other than trypsin resulted in identification of only a small

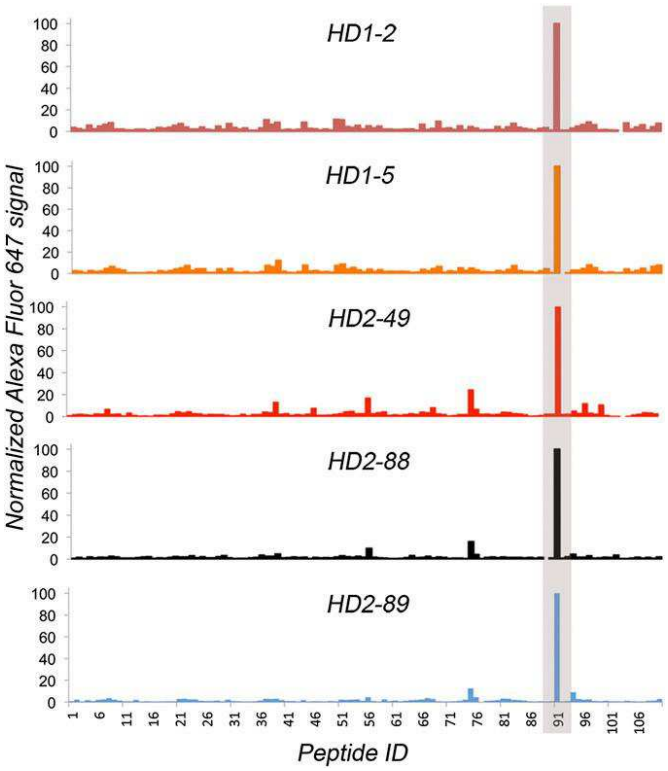

Fig. 5. Epitope mapping of rTT.C-specific antibodies. The histogram summarizes the normalized signal of five rTT.C-specific antibodies binding to a peptide array scan of an rTT.C sequence. The peptide array contained 110 15-mer peptides having a 12-residue overlap and spanning the entire rTT.C sequence. A negative control using Herceptin resulted in very low signal across the peptide array and positive control spots (absorbed full-length TT and rTT.C proteins). Computational modeling of three of these antibodies with the experimentally defined epitope is shown in Fig. S7.

number of additional low-frequency antibody clonotypes. Second, analysis of the V gene repertoire suggests that tryptic digestion suffices to generate peptide fragments of appropriate size for MS analysis for >92% of the V gene sequences in the 454 database (Fig. S3). Third, because of the presence of conserved trypsin cleavage sites adjacent to the CDR-H3/J region, the CDR-H3–derived peptides contain the highly conserved J region, and thus demonstrate similar characteristics and MS observability (Fig. S5). The similarity in the MS observability of CDR-H3–derived tryptic peptides therefore facilitates their relative quantitation by measuring peak intensities. Finally, with a detection limit of 0.1 nM, LC-MS/MS serum proteomics are easily able to identify antibodies that are present at physiologically relevant concentrations (i.e., concentrations $>K_d$), given that the theoretical ceiling of antibody affinity is 0.1 nM and the overwhelming majority of antibodies are expected to display higher K_d values, typically in the nanomolar range. For these reasons, the IgG-cd index provides a useful metric for evaluating differences in the breadth of the polyclonal vaccine response across samples.

Upon restimulation with antigen, mBCs undergo rapid expansion and differentiation into plasmablasts that emigrate into peripheral blood. In humans, the number of plasmablasts in blood reaches a peak value 7 d after immunization and then declines rapidly (32). Although the large majority of plasmablasts in the periphery undergo apoptosis, a small fraction, estimated at between 10% and 20% (30), are thought to be able to home into the BM, where they survive for extensive periods of time as LLPCs, producing the antibodies that constitute the long-term serological memory. Comparison of the plasmablast clonotypic repertoire with the anti-TT serum IgG repertoire revealed that ∼5% of the peak TT^+ plasmablast clonotypes contribute to the steady state (long-term) serological memory. These "successful" TT^+ plasmablast clonotypes were mostly present at high frequencies, cumulatively 38% and 30% (HD1 and HD2, respectively) of the total 454 reads in the day 7 TT^+ plasmablast population. Still, the data in Fig. 3 reveal that there are a significant number of the responding TT^+ plasmablast clonotypes that do not contribute significantly to the long-term serological memory, suggesting that these plasmablasts are mostly short-lived populations of B cells. The cellular features that render some highly expanded plasmablast clonotypes capable of producing antibody at a steady state, whereas others do not, will need to be defined further.

We observed a striking polarization of the serological repertoire, with a dominant antibody clonotype comprising 20% of the total peptide counts detected by LC-MS/MS, and estimated it to be present at a concentration >10 μg/mL in serum. Only four to seven antibody clonotypes in every sample analyzed were estimated to be present at concentrations >1 μg/mL, whereas the majority of the antibodies in the serum repertoire were detected at lower concentrations, generally within the range of 1–4 nM. The level of polarization seen here in the serum antibody repertoire, with a handful of serum antibody clonotypes dominating the response, displays some similarities to the peripheral B-cell repertoire determined by next-generation sequencing following flu immunization (33). Further, some serum antibodies (e.g., HD2-7) appear to be present in the serum at a concentration below the K_d value (Fig. 4, K_d for formaldehyde-treated TT =18.6 nM compared with an estimated serum concentration of 5 nM). Upon challenge with TT, such antibodies would be expected to be largely in the unbound state at equilibrium, and thus may not contribute significantly to protection.

Proteomic serum antibody profiling of the vaccine response at a molecular level facilitates a repertoire-wide analysis of neutralizing functionality of the constituent antibodies, which presents obvious utility in probing the efficacy of a vaccine in eliciting targeted responses toward neutralizing epitopes. For both donors, we find that recombinant IgGs encoded by the dominant clonotype in HD1 and HD2 bind to formaldehyde-treated TT with subnanomolar affinity; however, because they do not recognize the rTT.C fragment of the toxoid, they are unlikely to play an important role in neutralization despite their presence at high concentrations. Likewise, the majority (eight of 13) of serum antibodies tested also did not recognize rTT.C. In HD1, two of the top most abundant antibodies examined, HD1-2 and HD1-5 (observed at an estimated serum concentration of 10 μg/mL and 2 μg/mL, respectively) recognized native rTT.C protein with high affinity. Remarkably, epitope mapping revealed that both HD1-2 and HD1-5, as well as all serum anti-rTT.C IgGs identified in HD2, recognize the same linear epitope. Rosetta docking further showed that antibody binding to rTT.C occludes the ganglioside-binding site used by the toxin to gain entry into cells. These findings reveal that the existence of an immunogenic epitope at the ganglioside-binding site of the TT plays an important role in the elicitation of antibodies with neutralizing potential, and thus provides a molecular-level explanation of the action of the vaccine. Molecular-level identification and subsequent analysis of serum antibody-binding functionality provide obvious utility in understanding the effectiveness of a vaccine in eliciting neutralizing vs. nonneutralizing antibodies, as well as in understanding the effect of this balance on vaccine outcome (15).

Materials and Methods

Vaccination and Titers. A healthy male and female donor each received a booster vaccination comprising TT/DT [20 international units (IU) TT and 2 IU DT; Sanofi Pasteur MSD GmbH] after informed consent had been obtained. Serum anti-TT titers were determined in triplicate on ELISA plates coated with 2 μg/mL purified vaccine-grade TT (Statens Serum Institut). The anti-TT World Health Organization International Standard for Tetanus Immunoglobulin, Human (National Institute for Biological Standards and Control

code TE-3) was included in triplicate to allow conversion to international units per milliliter.

High-Throughput Sequencing of V_H and V_L Repertoires. Peripheral blood mononuclear cells were isolated and stained for FACS sorting as described in *SI Materials and Methods*. For each FACS-sorted B-cell population, first-strand cDNA was generated from total RNA using a SuperScript RT II kit (Invitrogen) and oligo-dT primer. V_λ, V_κ, and V_H repertoires were PCR-amplified as described (34). These amplified V gene repertoires from each sorted B-cell population were sequenced using high-throughput 454 GS-FLX sequencing (University of Texas at Austin and SeqWright). Raw 454 fasta files were submitted to the international ImMunoGeneTics database High V-Quest for V gene sequence alignment (35). Unique, full-length V_H gene sequences for each donor were assigned into clonotypes as described above.

Proteomic Analysis of the Serum Antibodies to TT in Human Donors. For each sample, IgG was purified from 7 to 9 mL of serum by protein G enrichment, followed by pepsin digestion to generate F(ab′)2 and Fc fragments. Antigen-specific F(ab′)2 fragments were isolated by TT-affinity column chromatography. Elution and flow-through fractions were trypsin-digested, and resulting peptides were fractionated and sequenced by nanoflow LC-electrospray MS/MS on an Orbitrap Velos Pro hybrid mass spectrometer (Thermo Scientific). The resulting spectra were searched against a custom protein sequence database as described (20) and detailed in *SI Materials and Methods*. The IgG-cd index represents the total number of serum IgG clonotypes detected in the serum of an individual at a distinct time point. Absolute peptide quantitation was done using isotopically labeled peptides spiked into MS-ready donor samples at different stoichiometric amounts and searched against the donor sequence database with isotopic labels included as dynamic modifications in the search.

Construction and Characterization of Recombinant IgG. V_H/V_L pairing of HD2 day 7 TT^+ plasmablasts was carried out as described (19). In addition, synthetic genes encoding the high-frequency, proteomically identified V_H sequences were paired with the V_L repertoire from day 7 TT^+ plasmablasts, expressed as Fabs in *Escherichia coli*, and screened for TT specificity by ELISA. V_H/V_L pairs exhibiting high ELISA signal were cloned and purified as IgG from HEK 293F cells. IgG affinities for vaccine-grade TT, rTT.C, and formaldehyde-inactivated rTT.C were determined by competitive ELISA.

Peptide Microarrays. Pepstar peptide microarrays (JPT Peptide Technologies GmbH) were constructed from overlapping 15-mer peptides (12-aa overlap) derived from rTT.C. Each microarray included three identical subarrays as technical replicates. Control rTT.C and TT spots were at separate locations in the array. The binding of purified recombinant antibodies to the peptide array was carried out according to the manufacturer's instructions, with modifications detailed in *SI Materials and Methods*.

Antibody Homology Modeling and Clustering Methods. Antibody V_H and V_L structures were determined by structural homology modeling using the Rosetta 3.5 suite of protein structure prediction and design software (36). Antibody homology models were then docked to rTT.C (Protein Data Bank ID code 1AF9) and analyzed as described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Dr. Scott Hunicke-Smith for assistance with next-generation sequencing, Constantine Chrysostomou for assistance in data analysis, Chhaya Das for recombinant IgG expression, Prof. Jeffrey J. Gray and Dr. Daisuke Kuroda for advice and assistance on the Rosetta modeling, Alexa Rodin for assistance with experiments, T. Kaiser and J. Kirsch for assistance with cell sorting, and Prof. Brent L. Iverson for useful discussion and comments. Funding for this work was provided by the Clayton Foundation (G.G.), by Welch Foundation Grant F1515 (to E. M. Marcotte), by the Defense Advanced Research Projects Agency (G.G. and A.D.E.), and by National Institutes of Health (NIH) Grants 5 RC1DA028779 (to G.G.) and GM 076536 (to E. M. Marcotte). J.J.L. was supported by a postdoctoral fellowship from the Cancer Prevention Research Institute of Texas. The LTQ Orbitrap Velos Pro instrument was purchased with support by the NIH Western Research Center of Excellent in Biodefense (NIH Grant 5U54AI057156) and the Texas Institute for Drug and Diagnostics Development (Grant TI-3D).

1. Germain RN (2010) Vaccines and the future of human immunology. *Immunity* 33(4): 441–450.
2. Pulendran B, Ahmed R (2011) Immunological mechanisms of vaccination. *Nat Immunol* 12(6):509–517.
3. Rappuoli R (2007) Bridging the knowledge gaps in vaccine design. *Nat Biotechnol* 25(12):1361–1366.
4. Sallusto F, Lanzavecchia A, Araki K, Ahmed R (2010) From vaccines to memory and back. *Immunity* 33(4):451–463.
5. Amanna IJ, Carlson NE, Slifka MK (2007) Duration of humoral immunity to common viral and vaccine antigens. *N Engl J Med* 357(19):1903–1915.
6. Frölich D, et al. (2010) Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. *J Immunol* 185(5):3103–3110.
7. Mei HE, et al. (2009) Blood-borne human plasma cells in steady state are derived from mucosal immune responses. *Blood* 113(11):2461–2469.
8. Reddy ST, et al. (2010) Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat Biotechnol* 28(9):965–969.
9. Benner R, Hijmans W, Haaijman JJ (1981) The bone marrow: The major source of serum immunoglobulins, but still a neglected site of antibody formation. *Clin Exp Immunol* 46(1):1–8.
10. Amanna IJ, Slifka MK (2010) Mechanisms that determine plasma cell lifespan and the duration of humoral immunity. *Immunol Rev* 236:125–138.
11. Manz RA, Hauser AE, Hiepe F, Radbruch A (2005) Maintenance of serum antibody levels. *Annu Rev Immunol* 23:367–386.
12. McHeyzer-Williams M, Okitsu S, Wang N, McHeyzer-Williams L (2012) Molecular programming of B cell memory. *Nat Rev Immunol* 12(1):24–34.
13. Pape KA, Taylor JJ, Maul RW, Gearhart PJ, Jenkins MK (2011) Different B cell populations mediate early and late memory during an endogenous immune response. *Science* 331(6021):1203–1207.
14. Radbruch A, et al. (2006) Competence and competition: The challenge of becoming a long-lived plasma cell. *Nat Rev Immunol* 6(10):741–750.
15. Khurana S, et al. (2013) Vaccine-induced anti-HA2 antibodies promote virus fusion and enhance influenza virus respiratory disease. *Sci Transl Med* 5(200):200ra214.
16. Crowe JE (2013) Universal flu vaccines: Primum non nocere. *Sci Transl Med* 5(200): 200fs234.
17. Wilson PC, Andrews SF (2012) Tools to therapeutically harness the human antibody response. *Nat Rev Immunol* 12(10):709–719.
18. Cheung WC, et al. (2012) A proteomics approach for the identification and cloning of monoclonal antibodies from serum. *Nat Biotechnol* 30(5):447–452.
19. DeKosky BJ, et al. (2013) High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* 31(2):166–169.
20. Wine Y, et al. (2013) Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. *Proc Natl Acad Sci USA* 110(8):2993–2998.
21. Ataro P, Mushatt D, Ahsan S (2011) Tetanus: A review. *South Med J* 104(8):613–617.
22. Franz B, May KF, Jr., Dranoff G, Wucherpfennig K (2011) Ex vivo characterization and isolation of rare memory B cells with antigen tetramers. *Blood* 118(2):348–357.
23. González-García I, Rodríguez-Bayona B, Mora-López F, Campos-Caro A, Brieva JA (2008) Increased survival is a selective feature of human circulating antigen-induced plasma cells synthesizing high-affinity antibodies. *Blood* 111(2):741–749.
24. Fotinou C, et al. (2001) The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin. *J Biol Chem* 276(34):32274–32281.
25. Plotkin SA (2010) Correlates of protection induced by vaccination. *Clin Vaccine Immunol* 17(7):1055–1065.
26. Schauer U, et al. (2003) Levels of antibodies specific to tetanus toxoid, Haemophilus influenzae type b, and pneumococcal capsular polysaccharide in healthy children and adults. *Clin Diagn Lab Immunol* 10(2):202–207.
27. Moody MA, et al. (2011) H3N2 influenza infection elicits more cross-reactive and less clonally expanded anti-hemagglutinin antibodies than influenza vaccination. *PLoS ONE* 6(10):e25797.
28. Poulsen TR, Jensen A, Haurum JS, Andersen PS (2011) Limits for antibody affinity maturation and repertoire diversification in hypervaccinated humans. *J Immunol* 187(8):4229–4235.
29. Purtha WE, Tedder TF, Johnson S, Bhattacharya D, Diamond MS (2011) Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. *J Exp Med* 208(13):2599–2606.
30. Höfer T, et al. (2006) Adaptation of humoral memory. *Immunol Rev* 211:295–302.
31. Georgiev IS, et al. (2013) Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. *Science* 340(6133):751–756.
32. Odendahl M, et al. (2005) Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. *Blood* 105(4): 1614–1621.
33. Jiang (2013) Lineage structure of the human antibody repertoire in response to influenza vaccination. *Sci Transl Med* 5(171):171ra19.
34. Ippolito GC, et al. (2012) Antibody repertoires in humanized NOD-scid-IL2Rγ(null) mice and human B cells reveals human-like diversification and tolerance checkpoints in the mouse. *PLoS ONE* 7(4):e35497.
35. Lefranc MP, et al. (1999) IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 27(1):209–212.
36. Kaufmann KW, Lemmon GH, Deluca SL, Sheehan JH, Meiler J (2010) Practically useful: What the Rosetta protein modeling suite can do for you. *Biochemistry* 49(14): 2987–2998.

Supporting Information

Lavinder et al. 10.1073/pnas.1317793111

SI Materials and Methods

Vaccination and Titers. A male healthy donor (HD) and female HD each received a booster vaccination comprising tetanus toxoid (TT)/diphtheria toxoid [DT; 20 international units (IU) TT and 2 IU DT; Sanofi Pasteur MSD GmbH] after informed consent had been obtained from the individual patients and after approval by the Local Ethics Committee of the Charité University Medicine Berlin. Forty milliliters of blood was collected before vaccination (day 0) and subsequently at day 7, 3 mo, and 9 mo postboost (Institutional Review Board Study 2012-08-0031). Thereof, 10 mL of peripheral blood was collected into a single Vacutainer K-EDTA collection tube (BD Biosciences). The additional 30 mL of peripheral blood was collected into three (3 × 10 mL) Vacutainer SST II serum tubes (BD Biosciences), with a resultant ∼15 mL of serum at each time point. Collection of peripheral blood mononuclear cells (PBMCs) from the K-EDTA blood was performed by density gradient centrifugation over Histopaque 1077 (Sigma) according to the manufacturer's protocol.

Serum anti-TT titers were determined using ELISA plates that were coated overnight at 4 °C with 2 μg/mL purified vaccine-grade TT (Statens Serum Institut) in 50 mM carbonate buffer (pH 9.6). Coated plates were washed three times in PBST (PBS with 0.1% Tween 20) and blocked with 2% (wt/vol) milk in PBS for 2 h at room temperature. Triplicates of serum samples were serially diluted twofold in 2% (wt/vol) milk in PBS (1:1,000–1:128,000 serum dilution factor) and then added to the blocked plates and incubated for 1 h at room temperature. On the same plate, serial dilutions of the anti-TT World Health Organization International Standard for Tetanus Immunoglobulin, Human (National Institute for Biological Standards and Control code TE-3) were included in triplicate to allow conversion to international units per milliliter. After binding, ELISA plates were washed three times with PBST and incubated with 50 μL of anti-human kappa light chain–HRP secondary antibody [1:2,500 in PBS and 2% (wt/vol) milk; Sigma] for 30 min at 25 °C. Plates were washed three times with PBST, and 50 μL of Ultra TMB substrate (Thermo Scientific) was then added to each well and incubated at 25 °C for 5 min. Reactions were stopped using an equal volume of 1 M H_2SO_4, and absorbance was read at 450 nm (BioTek).

Fluorescence-Activated Cell Sorting Analysis and Sorting of B-Cell Populations. For the two donors receiving the TT/DT booster vaccination, PBMCs were stained for 15 min in PBS/0.2% BSA at 4 °C in the dark using the following antibodies: anti-CD3–Pacific Blue (PacB, clone UCHT1; BD Biosciences), anti-CD14–PacB (clone M5E2; BD Biosciences), anti-CD19–phycoerythrin-cyanine 7 (PECy7; clone SJ25C1, BD Biosciences), anti-CD27–Cy5 (clone 2E4; BD Biosciences), anti-CD38–PE (clone HIT2; BD Biosciences), anti-CD20–Pacific Orange (clone HI47; Invitrogen), and anti-IgD–peridinin/chlorophyll/protein complex–Cy5.5 (clone L27; BD Biosciences). TT-specific B cells were identified by binding to TT labeled with digoxigenin (TT-Dig; Novartis Behring), washed in PBS/0.2% BSA, and bound by the secondary antibody anti-Dig–FITC (Roche Diagnostics GmbH). Specificity of the staining was confirmed each time by blocking with pure TT. DAPI (Molecular Probes) was added before cell sorting to exclude dead cells. The following B-cell populations were sorted using a FACSAria II cell sorter (Becton Dickinson): $CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-$ plasma cells (PCs), $CD3^-CD14^-CD19^+CD27^+CD20^+IgD^-$ memory B cells, and $CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-TT^+$ PCs (TT^+ PCs). Those PCs that did not stain as TT^+ (i.e., total PC-sorted fraction depleted of TT^+ PCs during TT labeling) are notated as TT-depleted (TT^{dep}). B-cell subpopulations were sorted and collected into PBS/0.2% BSA, centrifuged at 500 × *g* for 10 min, aspirated, and then resuspended in TRI Reagent solution (Life Technologies) and frozen at −80 °C. For HD2, approximately half of the day 7 TT^+ plasmablast-sorted cells were washed and cryopreserved in DMSO/10% (vol/vol) FCS for high-throughput V_H/V_L pairing (1).

Amplification of the V_H and V_L Repertoires from B Cells. Briefly, total RNA was isolated according to the RiboPure Kit protocol (Life Technologies). First-strand cDNA generation was performed with 500 ng of isolated total RNA using a SuperScript RT II kit (Invitrogen) and oligo-dT primer. After cDNA construction, PCR amplification was performed to amplify the V_λ, V_κ, and V_H genes separately using a primer set described by Ippolito et al. (2). PCR reactions were carried out using Taq polymerase with Thermpol reaction buffer (New England Biolabs) with the following cycling conditions: 92 °C denaturation for 3 min; 92 °C for 1 min, 50 °C for 1 min, and 72 °C for 1 min for four cycles; 92 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min for four cycles; 92 °C for 1 min, 63 °C for 1 min, and 72 °C for 1 min for 20 cycles; and a final extension at 72 °C for 7 min. PCR products were gel-purified on a 1% agarose gel before sequencing.

High-Throughput Sequencing of V_H and V_L Repertoires. V gene repertoires isolated from sorted B-cell populations of both vaccinated human donors were sequenced using high-throughput 454 (Roche Diagnostics GmbH) GS-FLX sequencing (University of Texas at Austin and SeqWright). Raw 454 fasta files were submitted to the international ImMunoGeneTics database (IMGT) High V-Quest for V gene sequence alignment and germ-line gene assignment (3). IMGT aligned sequences were then processed by a series of quality control filters to remove truncated sequences (as defined by not having aligned FR1 or FR4 Ig regions) and sequences that contained stop codons or ambiguous sequencing reads. The remaining sequences were subsequently grouped according to unique full-length V gene amino acid sequencing for generation of the sequence database used for proteomic bioinformatics.

Clonotype Indexing. Unique, full-length, heavy-chain V gene sequences for each donor were assigned into clonotypes by single-linkage hierarchical clustering. The algorithm starts by stochastically seeding an amino acid sequence from the donor-specific database and building an initial cluster from the entire donor-specific database if the following conditions are met: The complementarity determining region (CDR)-H3 amino acid sequence is ≥90% identical and the CDR-H3 is ±1 amino acid in length (accounting for differential IMGT sequence alignments occasionally seen at the CDR-H3/J region junction). Variations in length were penalized the same as a substitution. Upon completion of the first round of clustering, a new seed is chosen and unassigned sequences are tested for the above constraints. This iterative process continues until no unassigned sequences remain. IMGT was used to annotate V and J germ-line use, followed by manual curation of the IMGT calls. Manual curation was restricted to clusters that comprised >98% of a single V germ line (but not 100%) or >95% of a single J germ line (but not 100%). Visual inspection of the sequence alignments (to the assigned germ-line element) for each of these clusters allowed manual confirmation or rejection of the V and J germ-line

annotations. After manual curation, each cluster was set to contain full V gene sequences that have the same V and J assignment by IMGT.

Proteomic Analysis of the Serum Antibodies to TT in HDs. For each time point examined in the two HDs, 7–9 mL of serum was diluted fourfold with Protein G binding buffer (Pierce), filtered with a 0.22-μm Acrodisc syringe filter (Pall Life Sciences), and passed over a 4-mL Protein G Plus agarose (Pierce) affinity column in gravity mode. The diluted serum was recycled three times over the column, which was subsequently washed with 15 vol of PBS and eluted with 5 vol of 100 mM glycine·HCl (pH 2.7). The elution was immediately neutralized with 2 mL of 1 M Tris·HCl (pH 8.5). The purified IgG was dialyzed overnight at 4 °C into 20 mM sodium acetate (pH 4.5) and concentrated to 10 mg/mL using an Amicon Ultra-4 centrifugal filter unit (EMD Millipore). Total protein G-purified IgG was digested with 1 mL of immobilized pepsin resin (Pierce) per 10 mg of IgG in 20 mM sodium acetate. Pepsin digestion was allowed to proceed for 7 h, shaking vigorously at 37 °C. The digestion of the IgG into F(ab′)2 was monitored by SDS/PAGE to ensure >95% cleavage. After cleavage was complete, pepsin resin was spun down at 2000 × *g* and the F(ab′)2-containing supernatant was aspirated into a separate 1.5-mL Eppendorf tube. The resin was then washed (vortexed) with 1 vol of 2× PBS and wash supernatant-separated by spinning through an Ultrafree-MC centrifugal filter unit (EMD Millipore) at 4,000 × *g*, with a total of three resin washes. The original supernatant and three wash supernatants were combined and pH-adjusted to 7.

Affinity chromatography for the isolation of antigen-specific F(ab′)2 was carried out by first coupling 5 mg of vaccine-grade TT onto 0.25 g of dry N-hydroxysuccinimide (NHS)–activated agarose (Pierce) by overnight rotation at 4 °C. The coupled agarose beads were washed with PBS, and unreacted NHS groups were blocked with 1 M ethanolamine (pH 8.3) for 30 min at room temperature, washed with PBS, and packed into a 5-mL chromatography column (Pierce). The column was then washed with 5 vol of 100 mM glycine (pH 2.7) to elute nonspecifically bound (unconjugated) antigen and then with 5 vol of PBS to equilibrate. F(ab′)2 fragments were applied to the antigen affinity column in gravity mode, with the flow-through collected and reapplied to the column three times. The column was subsequently washed with 15 vol of PBS and 5 vol of ddH_20, and then eluted into Maxymum Recovery 1.5-mL tubes (Axygen Scientific) using 1-mL fractions of 20 mM HCl (pH 1.7). The flow-through, wash, and each 1-mL elution fraction (immediately neutralized with NaOH/Tris) were analyzed by indirect ELISA against vaccine-grade TT. Elution fractions showing an ELISA signal were combined and concentrated under vacuum to a volume of ∼0.5 mL, and the combined concentrated affinity column elution was desalted into ddH_20 using a 2-mL Zeba spin column (Pierce). The desalted elution was further concentrated under vacuum to ∼0.5 mg/mL in a Maxymum Recovery 1.5-mL tube.

The combined desalted elution and an aliquot of the flow-through from the antigen affinity chromatography were each denatured in 50% (vol/vol) 2,2,2-trifluoroethanol (TFE), 50 mM ammonium bicarbonate, and 10 mM DTT at 60 °C for 60 min. The denatured, reduced F(ab′)2 was alkylated by incubation with 32 mM iodoacetamide (Sigma) for 1 h at room temperature and then quenched by the addition of 20 mM DTT for 1 h at room temperature. Denatured, alkylated F(ab′)2 samples were diluted 10-fold into 50 mM sodium bicarbonate to reach a final TFE concentration of 5% vol/vol. Trypsin digestion was carried out by adding trypsin at a ratio of 1:35 trypsin/protein and incubated overnight at 37 °C. Digestion was halted by the addition of formic acid to a 1% final concentration, and the sample volume was reduced to ∼100 μL under vacuum.

Peptides derived from proteolytic digestion were subjected to chromatographic separation on a C18 Hypersep SpinTip (Thermo Scientific), washed three times with 0.1% TFA, and eluted with 60% (vol/vol) acetonitrile in 0.1% TFA (4 × 50 μL). The C18 elution was dried under vacuum to ∼5 μL and diluted to 50 μL in 5% (vol/vol) acetonitrile and 0.1% TFA. Peptides were separated on an Acclaim PepMap rapid separation liquid chromatography (RSLC) reverse-phase column (Dionex; Thermo Scientific) using a 5–40% (vol/vol) acetonitrile gradient over 250 min at 300 nL/min (Dionex UltiMate 3000 RSLCnano; Thermo Scientific). Eluting peptides were directly injected onto an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) using a nano-spray source. The LTQ Orbitrap Velos was operated in the data-dependent mode with parent ion scans (MS1) collected at a resolution of 60,000. Monoisotopic precursor selection and charge state screening were enabled. Ions with a charge >+1 were selected for collision-induced dissociation fragmentation spectrum acquisition (MS2) in the LTQ Orbitrap Velos, with a maximum of 20 MS2 scans per MS1. Dynamic exclusion was active with a 45-s exclusion time for ions selected more than twice in a 30-s window.

The resulting spectra were searched against a custom protein sequence database consisting of the donor-specific V_H and V_L protein sequences concatenated to a human full-protein–coding sequence database (Ensembl 64, longest sequence/gene) and common contaminants protein list (MaxQuant) using SEQUEST (Proteome Discoverer 1.3; Thermo Scientific). The search specified full-tryptic peptides, allowing up to two missed cleavages. A precursor mass tolerance of 5 ppm was used, with a fragment mass tolerance set to 0.5 Da. Carbamidomethylation of cysteine (iodoacetamide) was selected as a static modification, and oxidized methionine was selected as a dynamic modification. High-confidence peptide spectrum matches (PSMs) were filtered at a <1% false discovery rate (determined by Percolator, Proteome Discoverer 1.3; Thermo Scientific). Due to the prevalence of leucine/isoleucine substitutions in V gene peptides, all Ile/Leu sequence variants were considered as a single peptide group. For remaining cases where multiple high-confidence identifications were scored for the same spectrum, PSMs with the lowest posterior error probability were selected. To limit false identifications arising from the unusually high rate of homology among V gene peptides, we used an average mass deviation (AMD) filter of <1.5 ppm for all peptides, as described (4). Briefly, observed masses were recalibrated (5) and AMD was calculated for all peptides. Only peptides with an AMD <1.5 ppm were included in the final dataset. To increase confidence in the identification of the proteolytic peptides further, only peptides observed in all three injections, and with a minimum of five cumulative total spectral counts in at least one time point (sample), were considered. Based on the Poisson distribution of spectral counts detected across the three separate injections for each sample, the statistical likelihood of a peptide that passes these filters not being observed due to stochastic noise alone is <1.6%.

For peptide quantification, the extracted-ion chromatogram (XIC) area was calculated for all precursor ions by the Precursor Ions Area Detector in Proteome Discoverer). Proteome Discoverer does not combine XIC areas of charge variants at the peptide level, so multiple XIC areas are commonly observed for single peptides. To account for this, a total XIC area was calculated as the sum of all unique XIC areas for a given peptide, including modifications. To normalize peptide abundance measurements across samples, peptide frequencies were calculated as the total XIC area of a peptide divided by the sum of all peptide XIC areas in that sample.

Due to the high rate of degeneracy among V gene peptides, most identified sequences mapped to multiple V genes in our database. We therefore limited our analysis to "informative" peptides, defined as proteolytic fragmentation products of sufficient length and uniqueness to map to a single clonotype in the

indexed (see above) donor-specific V gene sequence database. Clonotype frequencies within each sample were calculated using only clonotype informative peptides derived from the CDR-H3 region of the antibody. Clonotyped CDR-H3 peptides were determined to be antigen-specific if their frequency in the elution was at least 10-fold greater than their frequency in the flow-through.

Absolute Peptide Quantitation by Isotopically Labeled Peptide Spike-In. Synthetic peptides were ordered from Pierce Biotechnology with either isotopically heavy lysine [$^{13}C(6)^{15}N(2)$, +8.014 Da] or arginine [$^{13}C(6)^{15}N(4)$, +10.008 Da] as the C-terminal residue. Lyophilized peptides were solubilized in 10 mM acetic acid at a concentration of 1 mg/mL. For MS analysis, individual peptides were diluted and combined at a concentration of 10 nM, followed by serial dilution in 5% (vol/vol) acetonitrile and 0.1% formic acid to desired concentrations. For spike-in experiments, the peptide mix was serially diluted to 10× stock concentration and then added to MS-ready samples at a 1:10 volumetric ratio to reach the final concentration. Samples were analyzed by liquid chromatography (LC) tandem MS (MS/MS), and the data were processed as described above, with the exception that the isotopically heavy K and R residues were included as dynamic modifications in the search.

Construction and Characterization of Recombinant IgG. V_H/V_L pairing of HD2 day 7 TT^+ plasmablasts was carried out as previously described (1). In addition, for both donors, synthetic genes encoding the high-frequency, proteomically identified V_H sequences (for which V_H/V_L pairs were not found by the method described by DeKosky et al. 1) were paired with the V_L repertoire from day 7 TT^+ plasmablasts, expressed as Fabs in *Escherichia coli*, and clones were screened for antigen specificity by ELISA on 96-well plates. Specifically, the day 7 TT^+ plasmablast V_L repertoire of each donor was PCR-amplified with human V_K and J_K primer sets as described above and cloned into pFAB-S, a Fab expression vector derived from pMAZ360 (6). For each donor, proteomically identified V_H genes were synthesized as gBlock Gene Fragments (Integrated DNA Technologies) and separately cloned into the pFAB-S vector containing the donor-specific V_L library. Each of the four libraries was electroporated into DH10B [F–*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L *nup*G λ–] and confirmed to be $>10^5$ in library size. Approximately 200 single colonies of each library were inoculated into 96-well plates containing terrific broth (TB) growth media + ampicillin (100 μg/mL) + 1% glucose, grown on a plate shaker at 30 °C overnight, and used to inoculate 96-well plates by transferring 5 μL of seed culture into 150 μL of TB + ampicillin (100 μg/mL) + 1% glucose. Cultures were grown on a plate shaker at 37 °C to ∼0.5 OD_{600}, and cells were pelleted by centrifugation in a swinging bucket rotor at 2,500 × *g* for 10 min at 4 °C (Beckman Coulter) and resuspended in an equal volume of TB + ampicillin (100 μg/mL) + 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for overnight expression, with shaking at 25 °C. ELISA screening of *E. coli* cell lysates was performed as described previously (7). The V_L genes of the antibodies that exhibited the highest signals by indirect ELISA against vaccine-grade TT were analyzed by Sanger sequencing (3730XL DNA Analyzers; Applied Biosystems). These V_H/V_L pairs were cloned separately into the pMAZ-IgH and pMAZ-IgL vectors (6), using Gibson Assembly Master Mix (New England Biolabs) and separately transformed into *E. coli* Jude-1 [*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* (F′ *proAB lacI*q*Z*Δ*M15* Tn*10* [Tetr])]. After sequence validation, 20 μg of each pMAZ vector cloned with the V_H and V_L gene was purified, DNA-sequenced, and cotransfected into HEK 293F cells following the Freestyle MAX expression system instructions (Invitrogen). HEK 293F cells were grown for 6 d after transfection; medium was harvested by centrifugation, and IgG was purified by a protein A agarose (Pierce) chromatography column.

IgG affinities for vaccine-grade TT, recombinant TT fragment C (rTT.C), and formaldehyde-inactivated rTT.C (f-rTT.C) were determined by competitive ELISA using different concentrations of IgG in a serial dilution of antigen, ranging from 50 to 0.02 nM in the presence of 2% (wt/vol) milk in PBS. The concentrations of IgG used were chosen based on the signal given in an initial indirect ELISA in which a dilution series of each IgG was analyzed. Each sample was incubated overnight at room temperature to equilibrate. Separately, plates were coated overnight at 4 °C with 10 μg/mL antigen (vaccine-grade TT, rTT.C, or f-rTT.C) in 50 mM carbonate buffer (pH 9.6), washed three times in PBST, and blocked with 2% (wt/vol) milk in PBS for 2 h at room temperature. Equilibrated samples were then added to the blocked ELISA plates and incubated for 1 h at room temperature. After binding, ELISA plates were washed three times with PBST and incubated with 50 μL of anti-human kappa-HRP secondary antibody [1:2,500 in PBS with 2% (wt/vol) milk; Sigma] for 30 min at 25 °C. Plates were washed three times with PBST; 50 μL of Ultra TMB substrate (Thermo Scientific) was then added to each well and incubated at 25 °C for 5 min. Reactions were stopped using an equal volume of 1 M H_2SO_4, and absorbance was read at 450 nm (BioTek).

For LC-MS/MS analysis of monoclonal recombinant IgG, ∼20 pmol of various purified recombinant antibodies was denatured, reduced, alkylated, and trypsinized, and peptides were purified over a C18 reverse-phase SpinTip as described above. A 10-fold dilution of the tryptic peptides from each antibody was injected onto the Orbitrap Velos and processed as described above for serum proteomics. Individual tryptic peptide spectral counts (derived from the monoclonal IgG) were compared and plotted to compare tryptic peptide observability across the full monoclonal IgG sequence.

rTT.C Expression and Purification. Plasmid pMOPAC12 (8) encoding a pelB leader, followed by the rTT.C gene and a C-terminal 6× His tag, was transformed into BL21 *E. coli*. For protein expression, cells were grown in TB media + ampicillin (100 μg/mL) to an OD_{600} of ∼0.6–0.8 and induced by addition of 1 mM IPTG in a shake flask at 250 rpm for 5 h at 25 °C. Cells were centrifuged at 8,500 × *g* for 10 min at 4 °C (Beckman Coulter), and the supernatant was discarded; the cell pellet was resuspended in lysis buffer [20 mM phosphate buffer, 300 mM sodium chloride (pH 7.5)] and subsequently lysed using a French press cell disruptor (Thermo Scientific). Lysates were centrifuged at 35,000 × *g* for 30 min at 4 °C and filtered with a 0.22-μm Acrodisc syringe filter (Pall Life Sciences). The rTT.C protein in the supernatant was purified by nickel-nitrilotriacetic acid affinity chromatography (Qiagen) according to the manufacturer's instructions. After washing with 10 column volumes of lysis buffer supplemented with 60 mM imidazole, rTT.C was eluted using lysis buffer supplemented with 300 mM imidazole. Protein concentration was determined using an ND-1000 spectrophotometer (Thermo Scientific), and purity was evaluated by SDS/PAGE analysis. Purified rTT.C was subjected to formaldehyde treatment for inactivation as previously described (9).

Peptide Microarrays. Pepstar peptide microarrays were provided by JPT Peptide Technologies GmbH. The peptides on the arrays represent a linear scan of 15-mer peptides derived from rTT.C that overlap by 12 amino acids. Each microarray included three identical subarrays as technical triplicates. As a positive control, rTT.C and whole TT were spotted on separate locations in the array. The binding of purified recombinant antibodies to the peptide array was carried out according to the manufacturer's instructions (www.jpt.com), with the following modifications: (*i*) Primary antibodies in TBS-Tween (TBST; 0.1% vol/vol) were incubated with the peptide microarray at a concentration of 20 μg/mL; (*ii*) peptide microarray slides were washed five times with TBST, incubated with Alexa Fluor 647-affinipure mouse anti-

human IgG (Jackson ImmunoResearch), and subsequently washed five times with TBST and then five times with ddH_2O; and (*iii*) slides were dried by applying a gentle stream of Argon on the microarray surface. Fluorescent signals were acquired with a GenePix 4000B scanner (Molecular Devices) at a resolution with a pixel size of 10 μm. Signals obtained were normalized and plotted to reflect the relative intensities of the fluorescence signals.

Antibody Homology Modeling and Clustering Methods. Antibody V_H and V_L structures were determined by structural homology modeling using the Rosetta 3.4 suite of protein structure prediction and design software. Antibody V_H and V_L sequences were matched to a database of antibody crystal structures. Each CDR and framework region of an antibody sequence was individually matched in sequence and length to a corresponding structural segment in the database (10). The matching antibody segments were then grafted into contiguous V_H and V_L models by superimposition of four backbone residues comprising each end of a CDR onto four overlapping backbone residues of an adjacent framework fragment. Terminal residues at each end of a CDR loop were then further modeled and optimized through superposition of residues on the terminus of the CDR onto the framework residues to achieve correct loop geometries. The entire modeled V_H/V_L pair was then minimized and subjected to side-chain rotamer repacking. The orientation of V_H and V_L chains was optimized through a local refinement docking step (11). The CDR-H3 was further modeled through kinematic loop modeling and refinement (12). A total of 2,000 models per antibody were generated, and the model exhibiting the best energy (in Rosetta energy units), as well as CDR-H3 structure, was selected per antibody docked to fragment C of the tetanus neurotoxin (Protein Data Bank ID code 1AF9). Docking of each antibody was restricted to samples within 10-Å translations and rotations within 15° of an initial orientation that was defined by the experimentally determined region of epitope binding. A total of 25,000 docking decoys were generated per antibody (13). Further decoys were generated by restricting the antibody to local samples within 5-Å translations and rotations within 10° of the initial orientation. Decoys were filtered for overall energy as well as for interface energy. Filtered models were clustered by backbone rmsd. In this way, the backbone Cα atoms of each model within a cluster are within a 3-Å rsmd of all other members of that cluster. Models representing the lowest energy member (by interface energy) and the largest cluster were selected as representative of overall antibody–antigen binding. These models were further optimized through local refinement docking for accurate measurement of solvent-accessible surface area buried by antibody–antigen interactions.

1. DeKosky BJ, et al. (2013) High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* 31(2):166–169.
2. Ippolito GC, et al. (2012) Antibody repertoires in humanized NOD-scid-IL2Rγ(null) mice and human B cells reveals human-like diversification and tolerance checkpoints in the mouse. *PLoS ONE* 7(4):e35497.
3. Lefranc MP, et al. (1999) IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 27(1):209–212.
4. Wine Y, et al. (2013) Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. *Proc Natl Acad Sci USA* 110(8):2993–2998.
5. Cox J, Michalski A, Mann M (2011) Software lock mass by two-dimensional minimization of peptide mass errors. *J Am Soc Mass Spectrom* 22(8):1373–1380.
6. Mazor Y, Van Blarcom T, Mabry R, Iverson BL, Georgiou G (2007) Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli. *Nat Biotechnol* 25(5):563–565.
7. Reddy ST, et al. (2010) Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat Biotechnol* 28(9):965–969.
8. Hayhurst A, et al. (2003) Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis. *J Immunol Methods* 276(1-2): 185–196.
9. Paliwal R, London E (1996) Comparison of the conformation, hydrophobicity, and model membrane interactions of diphtheria toxin to those of formaldehyde-treated toxin (diphtheria toxoid): Formaldehyde stabilization of the native conformation inhibits changes that allow membrane insertion. *Biochemistry* 35(7):2374–2379.
10. Sivasubramanian A, Sircar A, Chaudhury S, Gray JJ (2009) Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. *Proteins* 74(2):497–514.
11. Sircar A, Gray JJ (2010) SnugDock: Paratope structural optimization during antibody-antigen docking compensates for errors in antibody homology models. *PLOS Comput Biol* 6(1):e1000644.
12. Mandell DJ, Coutsias EA, Kortemme T (2009) Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat Methods* 6(8): 551–552.
13. Chaudhury S, et al. (2011) Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS ONE* 6(8):e22477.

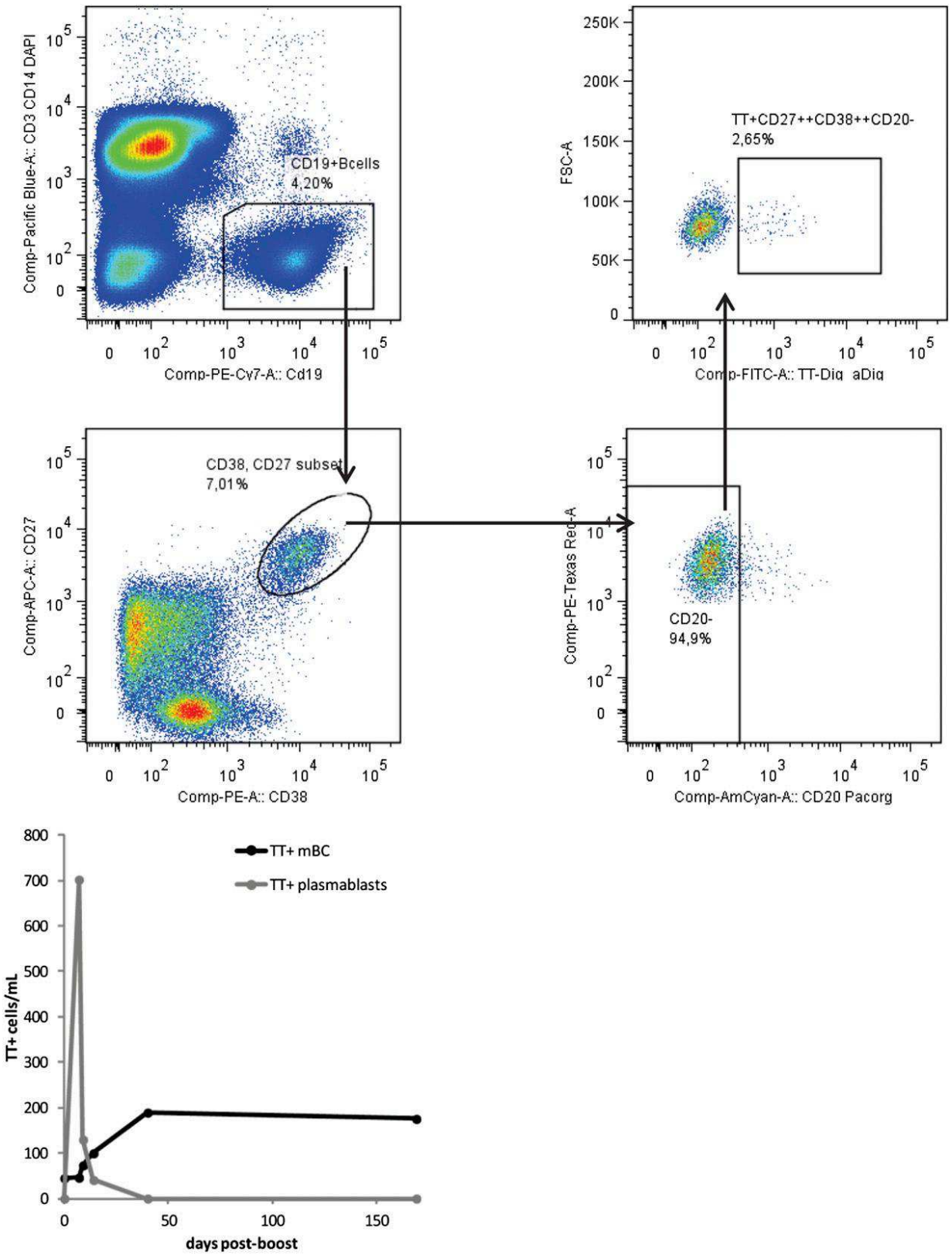

Fig. S1. Kinetics of TT^+ memory B cells (mBCs) and TT^+ plasmablasts in the peripheral blood of HD2. (*Upper*) Fluorescence-activated cell sorting (FACS) of $CD3^-CD14^-CD19^+CD27^{++}CD38^{++}CD20^-TT^+$ plasmablasts. Approximately 7% of day 7 B cells stained as $CD38^{++}CD27^{++}$ plasmablasts, and 3% of these stained as TT^+. (*Lower*) FACS TruCount (1) analyses of TT^+ mBCs and TT^+ plasmablasts on days 0, 7, 9, 14, 40, and 169 postboost. APC, allophycocyanin; Comp, compensated; Dig, digoxigenin; FSC, forward scatter; Pacorg, Pacific Orange.

1. Nicholson JKA, et al. (1997) Evaluation of a method for counting absolute numbers of cells with a flow cytometer. *Clin Diagn Lab Immunol* 4(3):309–313.

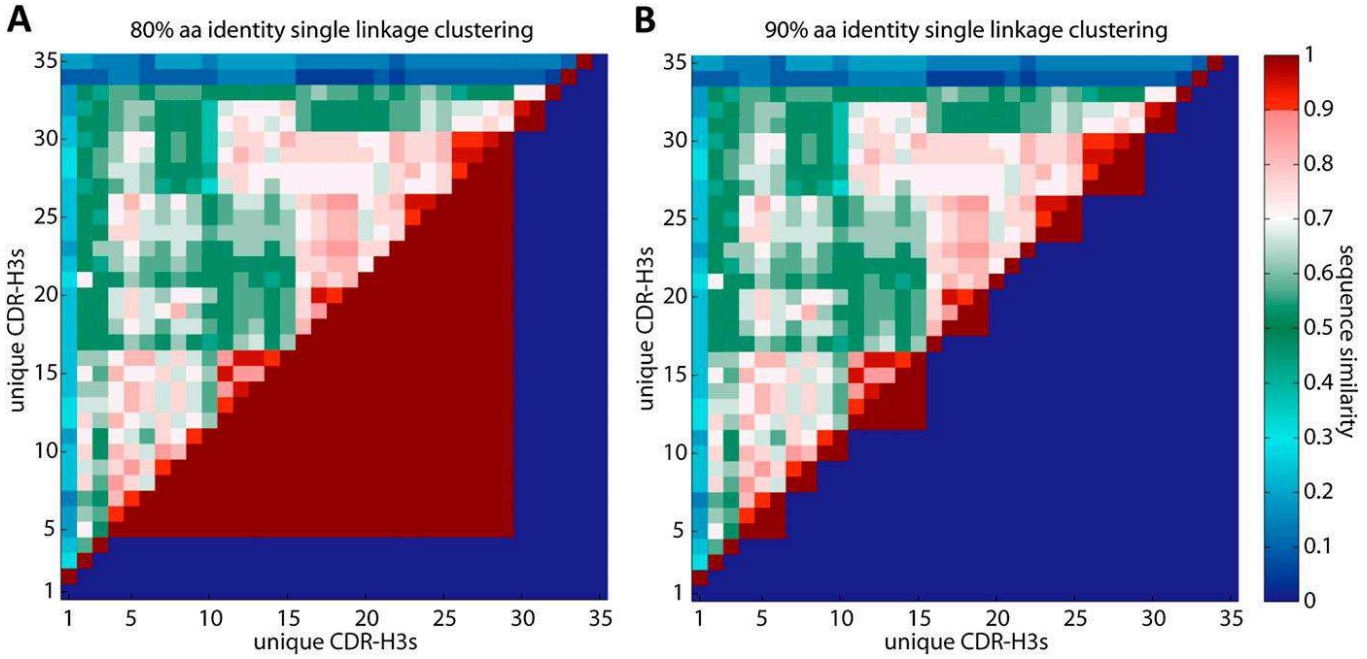

Fig. S2. Example set of sequences clustered using two different thresholds for amino acid identity in the CDR-H3. This example shows 35 sequences pulled from a larger test set. Half of the plot (upper left of the diagonal) shows the pairwise amino acid (aa) sequence identity among the 35 sequences, whereas the other half (lower right of the diagonal) shows the clonotype groupings (in red) formed by the single-linkage hierarchical clustering. (*A*) At an 80% identity threshold, the sequences mostly cluster into one clonotype in this example. However, this cluster (clonotype) has numerous CDR-H3s that are less than 60% identical (due to single linkage), clearly indicating the threshold is too low. (*B*) At a 90% identity threshold, the clonotypes formed represent groups of highly similar CDR-H3s.

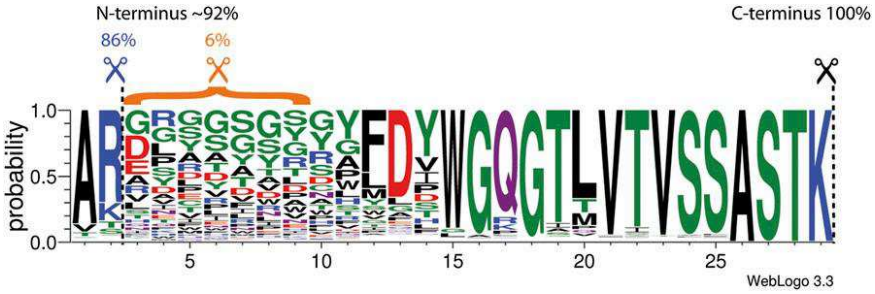

Fig. S3. Sequence logo of the CDR-H3/J region. The sites of trypsin cleavage for producing the CDR-H3–containing peptide are shown above in the form of a sequence logo produced from the donor-specific V gene sequence database (for the purposes of illustration here, HD2 14-residue CDR-H3). For all sequences within the DNA database, there exists a conserved trypsin cleavage site (R/K) in 86% of sequences. Of the remaining sequences, another 6% have an internal trypsin cleavage site in the CDR-H3. The C-terminal trypsin cleavage site resides within the constant region.

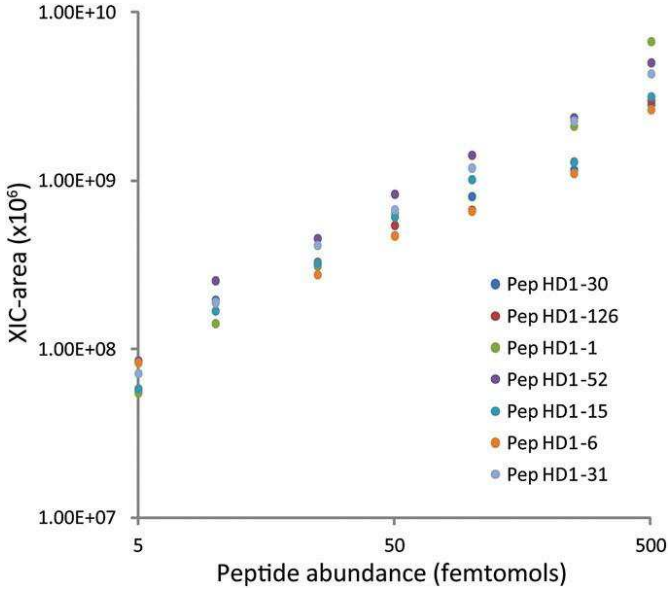

Fig. S4. Sensitivity of CDR-H3 peptide quantitation by precursor peak area. Isotopically labeled synthetic peptides corresponding to seven proteomically identified CDR-H3 sequences were spiked into HD1 samples at specific quantities and analyzed by LC-MS/MS as described in *SI Materials and Methods*, and peak areas were calculated from extracted precursor XIC areas. The peptide name is reflective of the donor and the TT^+ serum clonotype ranking (by postboost steady-state frequency) from which it is derived.

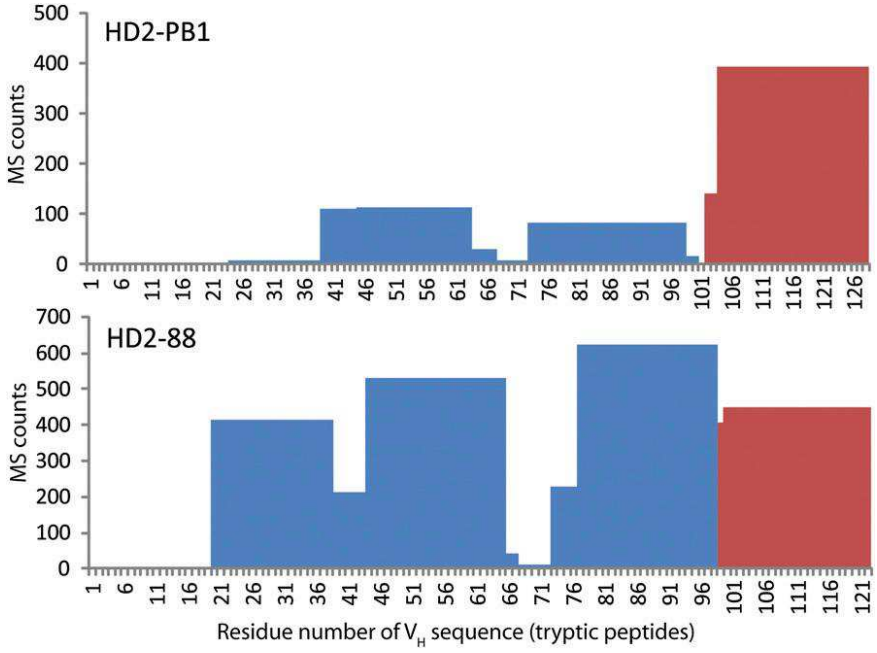

Fig. S5. Observed peptides from MS analysis of recombinant TT^+ mAbs. Purified recombinant anti-TT antibodies with V_H sequences identified either in the serum and TT^+ plasmablasts (HD2-88) or encoded in TT^+ plasmablasts only (HD2-PB1) were subjected to LC-MS/MS as described in *SI Materials and Methods*. CDR-H3 peptides for HD2-PB1 were not detected at any time point in the serum, despite the fact that this B-cell clonotype was identified in the TT^+ plasmablast compartment at a high frequency (9% of 454 sequencing reads) and was TT-specific (Fig. 4). The distribution of the antibody-derived tryptic peptides, in particular the clonotype-specific CDR-H3 peptides (red), indicates that the lack of observation of HD2-PB1 CDR-H3 peptides in serum-derived samples is not due to MS analytical artifacts.

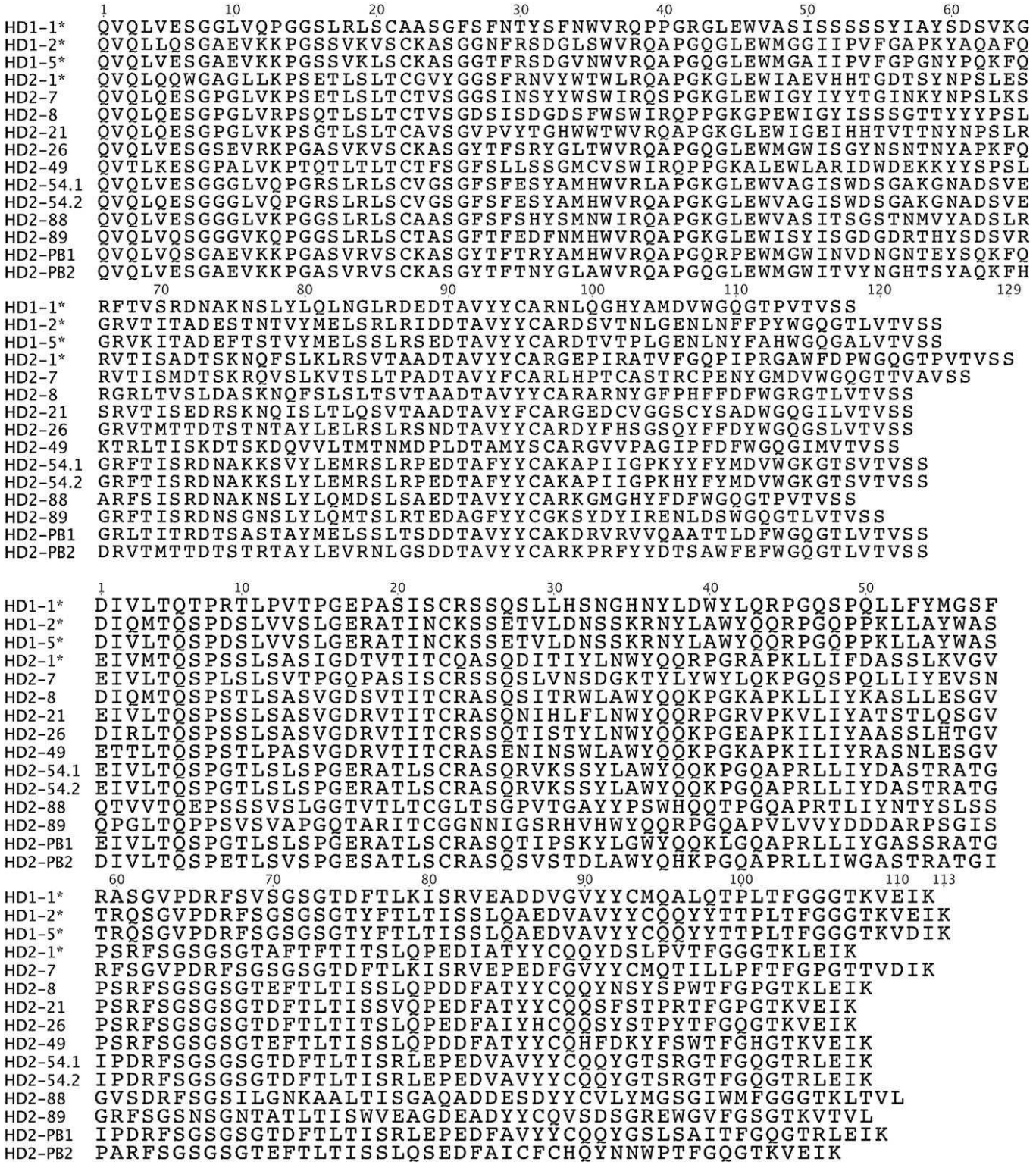

Fig. S6. Amino acid sequences of the V_H (*Upper*) and V_L (*Lower*) genes encoding the validated TT^+ antibodies. V_H/V_L pairings were determined by high-throughput single-cell pairing (1). The four sequences marked with an asterisk are the IgGs in which V_H/V_L pairing was discovered by screening of a library prepared by pairing a synthetic V_H with the V_L repertoire from day 7 TT^+ plasmablasts.

1. DeKosky BJ, et al. (2013) High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* 31(2):166–169.

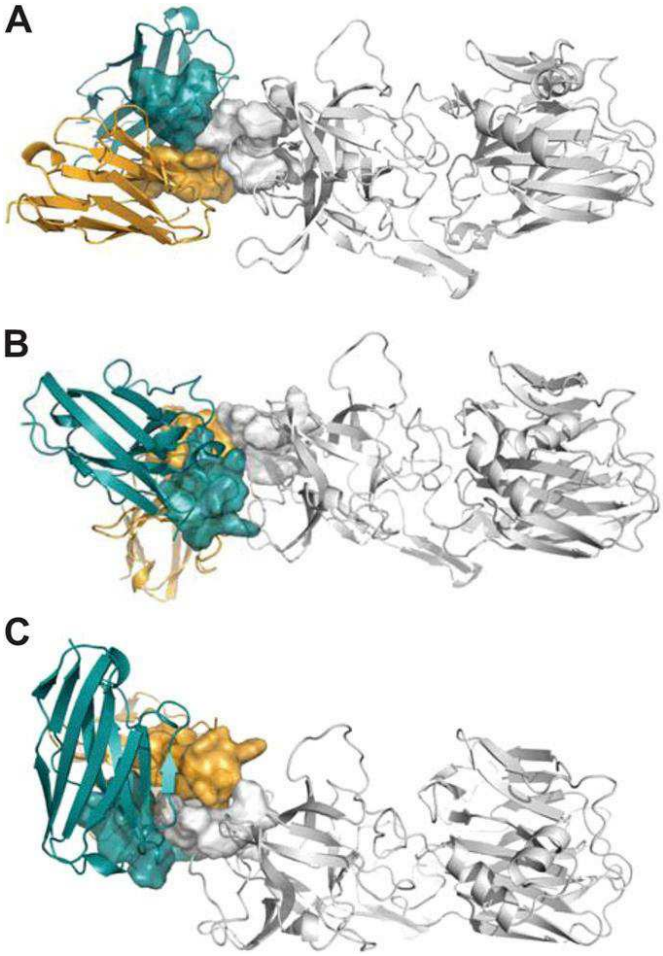

Fig. S7. Rosetta model of HD2-49 (*A*), HD2-88 (*B*), and HD2-89 (*C*) antibodies in complex with TT fragment C (TT.C). For all three antibodies, the V_H (cyan) and V_L (gold) chains were modeled and docked to the TT.C protein (gray). The *van der Waals* (*vdW*) surfaces for the CDR-H3 and CDR-L3 are shown, as is the *vdW* surface of the experimentally determined peptide epitope (Pep ID no. 91) on TT.C. The buried antibody solvent-accessible surface area (SASA) of the three models (HD2-49, HD2-88, and HD2-89) was 654.2 $Å^2$, 875 $Å^2$, and 892.8 $Å^2$, respectively.

Table S1. Summary of V_H 454 sequencing statistics and clonotype indexing

Cell population	Reads	Unique V genes	Unique clonotypes	Unique serum TT^+ clonotypes identified
HD1 TT^+ plasmablasts, day 7	16,051	6,898	922	25
HD1 TT^{dep} plasmablasts	24,456	9,750	1,804	10
HD1 mBCs, day 7	57,684	34,613	14,444	54
HD1 PBMCs, 3 mo	58,898	33,226	16,469	55
HD2 TT^+ plasmablasts, day 7	14,041	5,881	538	24
HD2 TT^{dep} plasmablasts	15,623	11,009	5,276	33
HD2 mBCs, day 7	12,693	9,047	5,672	25
HD2 PBMCs, 3 mo	27,969	12,818	5,254	24

TT^{dep}, TT depleted.

Table S2. Titers of HD1 and HD2 across five time points

Donor ID	Day 0	Day 7	Day 56	3 mo	9 mo
HD1*	2.0	2.3	4.7	4.5	4.5
HD2*	6.1	6.1	6.2	6.2	6.5

*Titers in international units per milliliter.